On the Babuška-Osborn approach to finite element analysis: L2 estimates for unstructured meshes

Article (Accepted Version)

ON THE BABUŠKA-OSBORN APPROACH TO FINITE ELEMENT ANALYSIS: L^2 ESTIMATES FOR UNSTRUCTURED MESHES

CHARALAMBOS G. MAKRIDAKIS

ABSTRACT. The standard approach to L^2 bounds uses the H^1 bound in combination to a duality argument, known as Nitsche's trick, to recover the optimal a priori order of the method. Although this approach makes perfect sense for quasi-uniform meshes, it does not provide the expected information for unstructured meshes since the final estimate involves the maximum mesh size. Babuška and Osborn, [1], addressed this issue for a one dimensional problem by introducing a technique based on mesh-dependent norms. The key idea was to see the bilinear form posed on two different spaces; equipped with the mesh dependent analogs of L^2 and H^2 and to show that the finite element space is inf-sup stable with respect to these norms.

Although this approach is readily extendable to multidimensional setting, the proof of the inf-sup stability with respect to mesh dependent norms is known only in very limited cases.

We establish the validity of the inf-sup condition for standard conforming finite element spaces of any polynomial degree under certain restrictions on the mesh variation which however permit unstructured non quasiuniform meshes. As a consequence we derive L^2 estimates for the finite element approximation via quasioptimal bounds and examine related stability properties of the elliptic projection.

1. INTRODUCTION

One of the fundamental, but nevertheless overlooked, open questions in the finite element analysis of elliptic problems is related to simple L^2 estimates for elliptic problems. The standard approach to L^2 bounds uses the H^1 bound in combination with a duality argument, known as Nitsche's trick, to recover the optimal a priori order of the method. Although this approach makes perfect sense for quasi-uniform meshes, it is less useful for unstructured meshes since the final estimate, say for piecewise linear elements, is of the form $\bar{h}\|hu\|_2$, where $\bar{h}$ denotes the maximum mesh size; see below for a precise notation. The typical estimates one finds in textbooks and many papers are of the form $\bar{h}^2\|u\|_2$. Now that questions related to the design and analysis of adaptive algorithms are relevant, there is a need to revise our point of view on this issue.

In a seminal paper published in 1980 Babuška and Osborn, [1], were the first to raise this question. This paper thoroughly analysed the one dimensional problem by introducing a technique based on mesh-dependent norms. The key idea was to see the bilinear form posed on two different spaces; equipped with the mesh dependent analogs of L^2 and H^2 and to show that the finite element space is inf-sup stable with respect to these norms. As a result, the stability of the elliptic projection as well as quasi-optimality results were established.

Although this approach is readily extendable to multidimensional setting, the proof of the inf-sup stability with respect to mesh dependent norms is known only in very limited cases. In this paper we establish the validity of the inf-sup condition for standard conforming finite element spaces of any polynomial degree under certain restrictions on the mesh variation which, however, permit unstructured non

---

Date: February 9, 2018.

Partially supported by the EU Horizon 2020 research and innovation programme under the Marie Skodowska-Curie project ModCompShock (`modcompshock.eu`) agreement No 642768.

quasiuniform meshes. As a consequence we derive L^2 estimates for the finite element approximation via quasi-optimal bounds and examine related stability properties of the elliptic projection.

2. Preliminaries

2.1. **Model problem and notation.** We denote the L^2-inner product by $(\cdot,\cdot) = (\cdot,\cdot)_\Omega$ on a domain $\Omega \subset \mathbb{R}^2$, and the corresponding L^2-norm $\|\cdot\|_\Omega$; the subscript Ω will be suppressed in most of the cases involving Ω, but it will remain in other domains $\mathcal{O}$ as $(\cdot,\cdot)_\mathcal{O}$ and $\|\cdot\|_\mathcal{O}$. The standard Sobolev spaces $W_p^m(\Omega)$, $H^m(\Omega) = W_2^m(\Omega)$ and $H_0^1(\Omega)$ will be used throughout. The corresponding norms and semi-norms will be denoted as usual by $\|\cdot\|_{m,p}$, $\|\cdot\|_m$ and $|\cdot|_{m,p}$, $|\cdot|_m$ respectively.

In $\Omega \subset \mathbb{R}^2$, we consider the boundary value problem:

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega, \tag{1}$$

for $f \in L^2(\Omega)$. The weak formulation reads,

$$B(u,v) = (\nabla u, \nabla v\,) = (f,v) \quad \text{for all } v \in H_0^1(\Omega). \tag{2}$$

We shall assume that standard elliptic regularity estimates are valid. For simplicity we consider convex polygonal domains Ω. We notice that the assumption $\Omega \subset \mathbb{R}^2$ is made only to simplify the presentation and it is not essential; similar results hold for $\Omega \subset \mathbb{R}^3$ with straightforward modifications.

Let $\mathcal{T}$ be a shape regular conforming subdivision of the domain Ω into elements K, [3]; the standard finite element space is denoted by

$$\mathbb{V}_h := \{v \in C(\overline{\Omega}) : v|_K \in \mathbb{P}_p(K),\ \ v|_{\partial\Omega} = 0\},$$

where $\mathbb{P}_p$ denotes the space of polynomials of degree at most p. We assume that to the underlined mesh it is associated an H^1 piecewise linear mesh-size function $h : \bar{\Omega} \to \mathbb{R}_+$, which is at each element K equivalent to $h_K = \operatorname{diam} K$. Such a mesh function can be constructed, e.g., by defining its value at a given vertex z to be the average of h_K for all K sharing z. Thus, for $v \in L^2(\Omega)$,

$$\|h\,v\| \quad \text{and} \quad \Big\{ \sum_{K\in\mathcal{T}} h_K^2 \|v\|_K^2 \Big\}^{1/2}, \tag{3}$$

are equivalent with constants independent of $\mathcal{T}$ but possibly depending on the shape regularity. Further, let $\Gamma = \cup_{K\in\mathcal{T}} \partial K \backslash \partial\Omega$ be the set of the inner faces of the decomposition.

Let e be a generic face of the decomposition. We shall use the notation

$$|v|_e^2 = \int_e v^2 dS\,. \tag{4}$$

Further, for a possibly discontinuous v on e, $[\![v]\!]_e$ denotes the jump of the traces of v on e .

2.2. **Mesh dependent norms.** Following standard notation introduced in [1, 2] we define,

$$H_h^2 := \{v \in H^1(\Omega) : v|_K \in H^2(K)\,\}, \tag{5}$$

and the mesh dependent 2-norm via

$$\|v\|_{2,h} = \Big\{ \|v\|_1^2 \ + \sum_{K\in\mathcal{T}} \|\Delta v\|_K^2 + \sum_{e\in\Gamma} h_e^{-1} \int_e [\![\partial_n u]\!]_e^2 \Big\}^{1/2}. \tag{6}$$

Note the inclusion of the $\|\cdot\|_1$ term to ensure that $\|\cdot\|_{2,h}$ is a norm. Further on H^1 we define

$$\|v\|_{0,h} = \Big\{ \ \|v\| + \sum_{e\in\Gamma} h_e |v|_e^2 \ \Big\}^{1/2}. \tag{7}$$

The motivation to consider these norms emanates by integrating by parts the weak formulation

$$B(u,v) = (\nabla u, \nabla v) = \sum_{K\in\mathcal{T}} -(\Delta u, v)_K + \sum_{e\in\Gamma} \int_e [\![\partial_n u]\!]_e \, v\,. \tag{8}$$

Then by assuming that $u \in H^2_h$ and $v \in H^1_0$ we have the natural continuity bound

$$|(\nabla u, \nabla v)| \le \|u\|_{2,h}\, \|v\|_{0,h} \tag{9}$$

and by reversing the roles of u and v we have

$$|(\nabla u, \nabla v)| \le \|v\|_{2,h}\, \|u\|_{0,h}\,. \tag{10}$$

Notice that the factors h_e and h_e^{-1} in the interface terms of the mesh dependent norms are natural choices to achieve the right scaling in the finite element space. Further, standard inverse estimates, [3, 4, 2], imply that $\|\cdot\|_{0,h}$ is equivalent to L^2 norm on $\mathbb{V}_h$:

$$c_0\, \|\chi\|_{0,h} \le \|\chi\| \le \|\chi\|_{0,h}\,, \qquad \forall \chi \in \mathbb{V}_h\,. \tag{11}$$

2.2.1. *The inf-sup condition.* Notice that both norms make sense in our finite element space $\mathbb{V}_h$. Our main goal is to prove the discrete *inf-sup* stability of the bilinear form $B(u,v)$ with respect to the mesh-dependent norms introduced above, that is:

(IS-$\mathbb{V}_h$) : *There exists a constant $\beta > 0$ independent of h such that for all $w \in \mathbb{V}_h$ there holds*

$$\sup_{v\in\mathbb{V}_h} \frac{B(w,v)}{\|v\|_{2,h}} \ge \beta\, \|w\|_{0,h}\,. \tag{12}$$

Let us remark here that we consider $B(w,v)$ on the *same* space $\mathbb{V}_h$ for both trial and test functions but equipped with two different norms. Notice that in the original paper [1] the analysis was relying on both the continuous and discrete inf-sup condition. However, for our purposes, the somewhat cumbersome continuous inf-sup condition is not needed.

Condition (12) along with the continuity of the bilinear form is enough to guarantee the stability of the finite element projection with respect to $\|\cdot\|_{0,h}$ norm and optimal approximation properties in L^2. To see this we introduce first the elliptic projection operator $R : H^1 \to \mathbb{V}_h$ by

$$(\nabla R\, v, \nabla \chi) = (\nabla v, \nabla \chi), \qquad \forall \chi \in \mathbb{V}_h\,. \tag{13}$$

Then (12) and (9) imply

$$\|R\,u\|_{0,h} \le \frac{1}{\beta} \sup_{v\in\mathbb{V}_h} \frac{B(R\,u,v)}{\|v\|_{2,h}} = \frac{1}{\beta} \sup_{v\in\mathbb{V}_h} \frac{B(u,v)}{\|v\|_{2,h}} \le \frac{1}{\beta}\, \|u\|_{0,h}\,. \tag{14}$$

The above stability bound is the closest we can get as far as the stability of the elliptic projection R in L^2 is concerned. In fact, a counterexample provided in [1] shows that the much desirable bound $\|R\,u\| \le C\, \|u\|$ is not true in general.

The main result of this note is

Theorem 2.1. *Assume that the mesh of the finite element space is such that the mesh function h has smooth enough variation, in the sense that $\|\nabla h\|_\infty \le \mu$ for μ small enough. Assume further that standard elliptic regularity estimates for the dual problem to* (1) *are valid. Then the inf-sup condition* **(IS-$\mathbb{V}_h$)** *holds.*

As it is typical in finite element analysis L^2 bounds require elliptic regularity estimates of the dual problem, which in our self adjoint case is the Laplace equation. This theorem will be proved in the next section. Among other consequences of this we mention the stability of the elliptic projection, the symmetric approximability bound (quasi-optimality) and the optimal L^2 estimate of the finite element solution. We summarise them in the following:

Theorem 2.2. *Under the assumptions of Theorem 2.1, we have the bounds*

$$\|R\,u\|_{0,h} \leq \frac{1}{\beta}\,\|u\|_{0,h}\ , \tag{15}$$

and

$$\|R\,u-u\|_{0,h} \leq (1+\frac{1}{\beta})\,\inf_{\chi\in\mathbb{V}_h}\,\|u-\chi\|_{0,h}\,. \tag{16}$$

As a consequence, if u *is the solution of* (2) *and* u_h *denotes its finite element approximation, there holds,*

$$\|u-u_h\| \leq C\,\|h^{p+1}D^{p+1}u\|\,, \tag{17}$$

where

$$\|h^sD^su\| := \Big\{\sum_{K\in\mathcal{T}} h_K^{2s}\,|v|_{s,K}^2\Big\}^{1/2}\,. \tag{18}$$

Theorem 2.2 follows directly by Theorem 2.1. Indeed, an immediate consequence of (15) is

$$\|R\,u-u\|_{0,h} \leq\ \|R\,(u-\chi)\|_{0,h}+\|u-\chi\|_{0,h} \leq (1+\frac{1}{\beta})\,\|u-\chi\|_{0,h}\,, \qquad \chi\in\mathbb{V}_h\,. \tag{19}$$

Hence the symmetric error bound (16) follows. Since $\|\cdot\|_{0,h}$ controls the L^2 norm, choosing now an appropriate interpolant Π, [3], such that

$$\|u-\Pi u\|_{0,h} \leq C\|h^{p+1}D^{p+1}u\|\,,$$

yields the desired bound. Since $p\geq 1$ one may choose the standard Lagrange interpolant. Then the above bound is a consequence of standard interpolation bounds along with scaled trace inequalities.

2.2.2. *Bibliographical remarks.* As far as the L^2 estimates for the finite element method are concerned, a standard duality argument implies the bound, [13, 3, 4],

$$\|u-u_h\| \leq C\,\bar{h}\|h^pD^{p+1}u\|\,, \tag{20}$$

where $\bar{h}=\max_{K\in\mathcal{T}}h_K$. This is in contrast to natural H^1 bounds derived by Cea's Lemma:

$$\|u-u_h\|_1 \leq C\,\|h^pD^{p+1}u\| = \Big\{\sum_{K\in\mathcal{T}} h_K^{2p}\,|v|_{p+1,K}^2\Big\}^{1/2}. \tag{21}$$

Notice that (21) holds without any restriction in the mesh. Although the nature of these bounds is similar for quasiuniform meshes, for unstructured decompositions this is obviously not the case.

In one space dimension the analysis in Babuška and Osborn [1] implies (17) without any restriction on the mesh. Still in one dimension, using a different approach, in Chapter 0 of Brenner and Scott [3], (17) was derived under the assumption $\|\nabla h\|_\infty \leq \mu$ for μ small enough. The case of piecewise linear elements in higher dimensions was treated in Eriksson and Johnson [7], see also [6], under the same assumption. Still under the same restriction on the mesh variation and for any polynomial degree, the analysis of Demlow and Stevenson [5] implies (17) but with the additional inclusion of a *data oscilation* term, i.e.,

$$\|u-u_h\| \leq C\,\{\|h^{p+1}D^{p+1}u\|+\|h^2(f-Pf)\|\,\}\,, \tag{22}$$

where Pf is an appropriate projection of the right hand side onto $\mathbb{V}_h$. In [5] the convergence of an adaptive algorithm in L^2 was demonstrated showing in addition that unstructured meshes satisfying the condition $\|\nabla h\|_\infty \leq \mu$ can be explicitly designed.

The inf-sup condition **(IS-**$\mathbb{V}_h$**)** has an independent interest that goes beyond its application to derive error estimates as (17) in Theorem 2.2. Both (15) and (19) are more general than (17) and they are important in their own right. These bounds and the condition **(IS-**$\mathbb{V}_h$**)** have various applications. Among others, they can be used in the analysis of mixed methods as in [2], in deriving quasioptimal bounds for time-dependent problems, [11], and in a posteriori analysis and the convergence of adaptive algorithms, [9], [12]. The appearance of $\|\cdot\|_{0,h}$ norm in the bounds of Theorem 2.2 looks natural in view of the results in [15]. As mentioned above it was derived in [1] for unstructured meshes in one dimension and in higher dimensions in [2] for quasiuniform meshes. The analysis in [7], combined with the framework used herein, can lead to the verification of **(IS-**$\mathbb{V}_h$**)** for piecewise linear elements under the assumption $\|\nabla h\|_\infty \leq \mu$.

As mentioned, we as well assume μ to be small enough. Theorem 2.1, provides a clear answer to inf-sup stability with respect to mesh dependent norms for conforming elements and any polynomial degree. Theorem 2.2 provides a natural extension of the standard H^1 bounds to L^2 when unstructured meshes are considered. It remains open whether the restriction on μ is really needed. As far as the method of proof is concerned, we emphasise that it is essentially a combination of some known techniques. The strategy to prove inf-sup stability by considering auxiliary problems is standard, and was used as well in [1, 2]. A new rather important feature is the use of an adaptation of arguments well known from the lower bounds of a posteriori analysis to define the functions w_1 and w_2, see Sections 3.2 and 3.3 below, which are crucial to show the various stability bounds of the auxiliary problem mentioned above. It is interesting to note that another instance of connection of a posteriori and a priori analysis is the work of Gudi [8] concerning a priori error estimates of the discontinuous Galerkin method. Lemma 3.1 was first used in [7] for $p = 1$, see also [6]. As far as we know the condition μ small was first used in the works of Eriksson and Johnson concerning a priori analysis for various time-dependent and stationary problems providing mathematical backup to adaptive computations, see e.g., [7, 6] and their references. Similar analytical tools were useful in the convergence of adaptive algorithms in L^2 in [5]. For completeness, we provide at the end of the paper a simple proof of Lemma 3.1. There we have used a super-approximation argument of [5] which has its origins in the fundamental paper of Nitsche and Schatz [14].

We have made a serious effort to present the analysis of this work as simply as possible. The arguments are quite clear and could serve as useful tools for improvements and possible extensions. The next section is devoted to the proof of Theorem 2.1.

3. Proof of Theorem 2.1

3.1. **Stability of an auxiliary problem.** The starting point of the proof is a standard dual problem. Consider $\rho \in \mathbb{V}_h$ given and fixed. Define $\Phi \in \mathbb{V}_h$ as the solution of

$$(\nabla \Phi, \nabla \psi) = (\rho, \psi) \quad \text{for all } \psi \in \mathbb{V}_h. \tag{23}$$

Our aim is then to show the stability bound

$$\|\Phi\|_{2,h} \leq c\,\|\rho\|\,. \tag{24}$$

Then (24) implies **(IS-**$\mathbb{V}_h$**)** . Indeed, since

$$(\nabla \Phi, \nabla \rho) = \|\rho\|^2,$$

(24) implies

$$(\nabla \Phi, \nabla \rho) \geq c \|\rho\| \|\Phi\|_{2,h}.$$

Or

$$\sup_{v \in \mathbb{V}_h} \frac{B(v, \rho)}{\|v\|_{2,h}} \geq \frac{(\nabla \Phi, \nabla \rho)}{\|\Phi\|_{2,h}} \geq c \|\rho\|,$$

and **(IS-$\mathbb{V}_h$)** follows in view of (11).

To prove (24) we need to control the three parts of the norm $\|\Phi\|_{2,h}$. Clearly, $\|\Phi\|_1 \leq C \|\rho\|$ in view of the Poincaré-Friedrichs inequality. To provide a clear description of the argument, we divide the rest of the proof in two steps: in Step 1 we control the terms involving $\sum_{K\in\mathcal{T}} \|\Delta\Phi\|_K^2$ and in Step 2 we control the jump terms $\sum_{e\in\Gamma} h_e^{-1} \int_e [\![\partial_n \Phi]\!]_e^2$.

3.2. **Step 1.** Define a w_1 which is a discrete object but not member of the space $\mathbb{V}_h$ such that

$$c_1 \sum_{K\in\mathcal{T}} \|\Delta\Phi\|_K^2 \leq (\nabla \Phi, \nabla w_1).$$

Then $R\, w_1 \in \mathbb{V}_h$ and (23) yields

$$\begin{aligned} c_1 \sum_{K\in\mathcal{T}} \|\Delta\Phi\|_K^2 &\leq (\nabla \Phi, \nabla w_1) = (\nabla \Phi, \nabla R\, w_1) \\ &= (\rho, R\, w_1). \end{aligned} \tag{25}$$

We will then show the following bounds

$$\|h\nabla w_1\| + \|w_1\| \leq c \Big(\sum_{K\in\mathcal{T}} \|\Delta\Phi\|_K^2 \Big)^{1/2} \tag{26}$$

and

$$\|R\, w_1\| \leq c \|h\nabla w_1\| + \|w_1\|. \tag{27}$$

Which combined with (25) yield the desired bound.

3.2.1. *Definition of w_1 and proof of* (26). We now define w_1 elementwise as

$$w_1|_K = -b_K \Delta\Phi|_K, \tag{28}$$

where b_K is the standard bubble function used in a posteriori analysis, [16]; for notational consistency, here and below we follow the notation in [10, Theorem 3.2]. Scaling and inverse-type arguments involving b_K are standard, but we repeat some of them for readers not familiar with a posteriori analysis. In particular, $b_K \in \mathbb{P}_3$ is positive on the interior of K and $\|b_K\|_\infty = 1$.

Then $w_1 \in H_0^1(\Omega)$, $w_1|_e = 0$ for all $e \in \Gamma$, and hence,

$$\begin{aligned} (\nabla \Phi, \nabla w_1) &= -\sum_{K\in\mathcal{T}} (\Delta \Phi, w_1) \\ &= \sum_{K\in\mathcal{T}} \int_K b_K |\Delta\Phi|^2. \end{aligned}$$

Now, since the norms $\|b_K^{1/2} \cdot \|_K$ and $\| \cdot \|_K$ on $\mathbb{P}_p$ are equivalent we have

$$(\nabla \Phi, \nabla w_1) \geq c_1 \sum_{K\in\mathcal{T}} \|\Delta\Phi\|_K^2,$$

where c_1 is a positive constant which possibly depends on p but it is independent of h. Further, using local inverse inequalities on $\mathbb{P}_{p+1}$, we obtain,

$$\|h\nabla w_1\| + \|w_1\| \leq\; c\,\|w_1\| \leq\; c\Big(\sum_{K\in\mathcal{T}} \|\Delta\Phi\|_K^2\Big)^{1/2}$$

and (26) follows.

3.2.2. *Proof of* (27). We use a variation of a duality argument used in [7]. Let g be the solution of the problem

$$-\Delta g = R\,w_1 \quad \text{in } \Omega, \qquad g = 0 \quad \text{on } \partial\Omega\,. \tag{29}$$

Then

$$\begin{aligned}\|R\,w_1\|^2 &= (\nabla\, R\,w_1, \nabla g\,)\\ &= (\nabla\,(R\,w_1 - w_1), \nabla g\,) + (\nabla\, w_1, \nabla g\,)\\ &=: I_1 + I_2\,.\end{aligned}$$

For I_2 we have, using once more that $w_1 \in H_0^1(\Omega), w_1|_e = 0$ for all $e\in\Gamma$,

$$I_2 = (\nabla\, w_1, \nabla g\,) = -\sum_{K\in\mathcal{T}}\int_K w_1\Delta g = (w_1, R\,w_1)\,.$$

Thus,

$$|I_2| \leq \|w_1\|\,\|R\,w_1\|\,.$$

To estimate I_1 we use Galerkin orthogonality, an appropriate interpolant $\Pi g \in \mathbb{V}_h$, and the regularity of the dual problem, to get

$$\begin{aligned}|I_1| &= |(\nabla\,(R\,w_1 - w_1), \nabla(g - \Pi g)\,)| \leq \|h\,\nabla\,(R\,w_1 - w_1)\|\,\|h^{-1}\nabla(g-\Pi g)\|\\ &\leq C\|h\,\nabla\,(R\,w_1 - w_1)\|\,|g|_{2,\Omega}\\ &\leq C\|h\,\nabla\,(R\,w_1 - w_1)\|\,\|R\,w_1\|\,.\end{aligned}$$

Combining the above estimates for I_1 and I_2 we get,

$$\|R\,w_1\| \leq C\|h\,\nabla\,(R\,w_1 - w_1)\| \;+ \|w_1\|\,.$$

Next, we shall use the following result regarding the elliptic projection operator

Lemma 3.1. *Let $\|\nabla h\|_\infty \leq \mu\,,\, \mu < 1\,,$ and $w\in H_0^1(\Omega)$. Then there exists a constant C such that*

$$\|h\,\nabla R\,w\| \leq C(\|h\,\nabla\,w\| \;+ \mu\,\|R\,w\|)\,.$$

We postpone the proof to the end of this section.

It is clear now that using Lemma 3.1 for $w = w_1$ in the above estimates and assuming that μ is small enough we conclude the proof of (27) and of Step 1, establishing that

$$\Big(\sum_{K\in\mathcal{T}} \|\Delta\Phi\|_K^2\Big)^{1/2} \leq C\|\rho\|\,. \tag{30}$$

3.3. **Step 2.** To control the jumps we shall use again a discrete function whose definition is motivated by a posteriori analysis. As already mentioned, we follow the notation of [10, Theorem 3.2].

Define w_2 as follows: Let e be an internal edge. Let b_e the bubble function of the side e which vanishes on its vertices. As before, we require b_e to be positive on the interior of e and $\|b_e\|_{\infty,e} = 1$. If K and K_e are the elements sharing e, define the discrete operator $\mathcal{J}_e : \mathbb{P}_{p-1}(e) \to H_0^1(\Omega)$ as follows:

(i) : $\operatorname{supp}\mathcal{J}_e(v_e) = K \cup K_e$

(ii) : $\mathcal{J}_e(v_e)|_e = b_e\, v_e$

(iii) : $\mathcal{J}_e(v_e)$ is extended to K and K_e by requiring $\mathcal{J}_e(v_e)$ to be linear along lines normal to e and being zero on $\partial K \cup \partial K_e \backslash e$. $\mathcal{J}_e(v_e)$ is set to zero outside $K \cup K_e$.

Define now the space

$$\mathbb{J}_h = \Big\{ w = \sum_{e\in\Gamma} \mathcal{J}_e(v_e) : \quad v_e \in \mathbb{P}_{p-1}(e),\ e \in \Gamma \Big\} . \tag{31}$$

It is clear that $\mathbb{J}_h$ is a finite dimensional subspace of $H_0^1(\Omega)$. We are ready to define now w_2. Let

$$w_e = \mathcal{J}_e\Big(\frac{1}{h_e}[\![\partial_n \Phi]\!]_e\Big) , \qquad e \in \Gamma , \tag{32}$$

and

$$w_2 = \sum_{e\in\Gamma} w_e \ . \tag{33}$$

Clearly, $w_2 \in \mathbb{J}_h$, and the restriction of w_2 to each element K is a polynomial. Further since $w_2 \in H_0^1(\Omega)$, we have,

$$\begin{aligned}(\nabla \Phi, \nabla w_2\,) &= -\sum_{K\in\mathcal{T}} (\Delta\, \Phi, w_2\,) + \sum_{e\in\Gamma} \int_e [\![\partial_n \Phi]\!]_e w_2 \\ &= -\sum_{K\in\mathcal{T}} (\Delta\, \Phi, w_2\,) + \sum_{e\in\Gamma} \frac{1}{h_e} \int_e b_e\, [\![\partial_n \Phi]\!]_e^2 ,\end{aligned}$$

where we used the fact that $w_2|_e = w_e$. As in the proof of Step 1, using the equivalence of $\|b_e^{1/2} \cdot \|_e$ and $\| \cdot \|_e$ on $\mathbb{P}_{p-1}(e)$ we have for some $c_2 > 0$,

$$\begin{aligned} c_2 \sum_{e\in\Gamma} \frac{1}{h_e} \int_e [\![\partial_n \Phi]\!]_e^2 &\le \sum_{e\in\Gamma} \frac{1}{h_e} \int_e b_e\, [\![\partial_n \Phi]\!]_e^2 \\ &= (\nabla \Phi, \nabla w_2\,) + \sum_{K\in\mathcal{T}} (\Delta\, \Phi, w_2\,) \\ &= (\nabla \Phi, \nabla R\, w_2\,) + \sum_{K\in\mathcal{T}} (\Delta\, \Phi, w_2\,) \\ &= (\rho, R\, w_2\,) + \sum_{K\in\mathcal{T}} (\Delta\, \Phi, w_2\,) \\ &\le \|\rho\|\, \|R\, w_2\,\| + \Big(\sum_{K\in\mathcal{T}} \|\Delta\, \Phi\|_K \Big)^{1/2} \|w_2\| \\ &\le C\, \|\rho\| \, \Big(\|R\, w_2\,\| + \|w_2\| \Big) , \end{aligned} \tag{34}$$

where in the last inequality we used the bound (30) proved in Step 1.

By construction w_2 belongs to the discrete space $\mathbb{J}_h$. It is a simple matter to check that

$$\|v\|_{0,\Gamma} := \Big(\sum_{e\in\Gamma} h_e \int_e |v|^2 \Big)^{1/2}, \tag{35}$$

is a norm in $\mathbb{J}_h$. Clearly $\|\cdot\|_{0,\Gamma}$ satisfies all the properties of a seminorm. Further, if $\|v\|_{0,\Gamma} = 0$ for $v \in \mathbb{J}_h$, then $v|_e = 0$ for all e and thus by definition of $\mathbb{J}_h$ $v = 0$ everywhere. Then by standard inverse and scaling arguments one gets,

$$\|h\nabla v\| + \|v\| \le C\,\|v\|_{0,\Gamma} \quad \text{for all} \quad v \in \mathbb{J}_h \,. \tag{36}$$

Assume for a moment that (27) holds for w_2 as well, that is,

$$\|R\,w_2\| \le c\left(\|h\nabla w_2\| + \|w_2\|\right). \tag{37}$$

Then, (37), (36) and (34) imply

$$\begin{aligned}
c_2 \sum_{e\in\Gamma} \frac{1}{h_e} \int_e [\![\partial_n \Phi]\!]_e^2 &\le C\,\|\rho\| \Big(\|R\,w_2\| + \|w_2\| \Big) \\
&\le C\,\|\rho\| \Big(\|h\nabla w_2\| + \|w_2\| \Big) \\
&\le C\,\|\rho\|\,\|w_2\|_{0,\Gamma} \\
&\le C\,\|\rho\| \Big(\sum_{e\in\Gamma} \frac{1}{h_e} \int_e b_e^2\,[\![\partial_n \Phi]\!]_e^2 \Big)^{1/2} \\
&\le C\,\|\rho\| \Big(\sum_{e\in\Gamma} \frac{1}{h_e} \int_e [\![\partial_n \Phi]\!]_e^2 \Big)^{1/2}.
\end{aligned}$$

Hence,

$$\Big(\sum_{e\in\Gamma} \frac{1}{h_e} \int_e [\![\partial_n \Phi]\!]_e^2 \Big)^{1/2} \le c\,\|\rho\|\,. \tag{38}$$

To complete the proof of Step 2 it remains therefore to verify (37).

3.3.1. *Proof of* (37). We proceed as in the proof of (27). Let g be the solution of the problem

$$-\Delta g = R\,w_2 \quad \text{in } \Omega, \qquad g = 0 \quad \text{on } \partial\Omega, \tag{39}$$

Then

$$\begin{aligned}
\|R\,w_2\|^2 &= (\nabla\,R\,w_2, \nabla g\,) \\
&= (\nabla\,(R\,w_2 - w_2), \nabla g\,) + (\nabla\,w_2, \nabla g\,) \\
&=: I_1 + I_2
\end{aligned}$$

Now, since w_2 is continuous at the interfaces, we obtain as before

$$I_2 = (\nabla\,w_2, \nabla g\,) = -\sum_{K\in\mathcal{T}} \int_K w_2 \Delta g = (w_2, R\,w_2) \le \|w_2\|\,\|R\,w_2\|\,.$$

The term I_1 is handled as in the proof of (27) and the proof is complete under the assumptions that Lemma 3.1 holds and μ is small enough.

It remains therefore to complete the

Proof of Lemma 3.1. The proof is quite simple, compare to [6, 7] for earlier similar results. A simple computation reveals,

$$\begin{aligned}
\|h\nabla R\,w\|^2 &= (h^2\nabla\,R\,w, \nabla\,R\,w) = (\nabla\,R\,w, \nabla\left(h^2\,R\,w\right)) - (\nabla\,R\,w, 2(h\nabla h)\,R\,w)\\
&= (\nabla\,(R\,w-w)\,, \nabla\left(h^2\,R\,w\right)) + (\nabla\,w, \nabla\left(h^2\,R\,w\right)) - (\nabla\,R\,w, 2(h\nabla h)\,R\,w)\\
&= (\nabla\,(R\,w-w)\,, \nabla\left(h^2\,R\,w\right)) + (\nabla\,w,\, 2(h\nabla h)\,R\,w))\\
&\quad + (\nabla\,w,\, h^2\,\nabla R\,w)\,) - (\nabla\,R\,w, 2(h\nabla h)\,R\,w)\\
&=: (\nabla\,(R\,w-w)\,, \nabla\left(h^2\,R\,w\right)) + Z\,.
\end{aligned}$$

Hence

$$|Z| \le 2\|h\nabla\,w\|\mu\,\|R\,w\| + \|h\nabla\,w\|\,\|h\nabla R\,w\| + 2\|h\nabla R\,w\|\mu\,\|R\,w\|\,.$$

Next we notice

$$(\nabla\,(R\,w-w)\,, \nabla\left(h^2\,R\,w\right)) = (\nabla\,(R\,w-w)\,, \nabla\left(h^2\,R\,w - \Pi[h^2\,R\,w]\right))\,.$$

Since $h|_K \in \mathbb{P}_1$, $|h^2|_{s,K} = 0$, for $s > 2$, and $|h^2|_{2,\infty} = 2|\nabla h|^2_{\infty,K}$, we have, compare to [5, p. 194], [14, p. 942],

$$\begin{aligned}
\|\nabla\left(h^2\,R\,w{-}\Pi[h^2\,R\,w]\right)\|_K &\le C\,h_K^p|h^2\,R\,w|_{p+1,K}\\
&\le C\,h_K^p\big(\|h^2\|_{\infty,K}|\,R\,w|_{p+1,K} + |\nabla(h^2)|_{\infty,K}|\,R\,w|_{p,K} + |\nabla h|^2_{\infty,K}|\,R\,w|_{p-1,K}\big)\\
&= C\,h_K^p\big(|h|_{\infty,K}|\nabla h|_{\infty,K}|\,R\,w|_{p,K} + |\nabla h|^2_{\infty,K}|\,R\,w|_{p-1,K}\big)\,.
\end{aligned}$$

Standard inverse inequalities imply,

$$\|\nabla\left(h^2\,R\,w{-}\Pi[h^2\,R\,w]\right)\|_K \le Ch_K\,(\mu+\mu^2)\,\|R\,w\|_K\,.$$

Hence, by the properties of the mesh function and since $\mu \le 1$,

$$\begin{aligned}
(\nabla\,(R\,w-w)\,, \nabla\left(h^2\,R\,w\right)) &= (\nabla\,(R\,w-w)\,, \nabla\left(h^2\,R\,w - \Pi[h^2\,R\,w]\right))\\
&\le C\sum_{K\in\mathcal{T}} h_K\,(\mu+\mu^2)\,(\|\nabla R\,w\|_K + \|\nabla w\|_K)\|R\,w\|_K\\
&\le C\sum_{K\in\mathcal{T}} (\mu+\mu^2)\,(\|h\nabla R\,w\|_K + \|h\nabla w\|_K)\|R\,w\|_K\\
&\le C\,\mu\,(\|h\nabla R\,w\| + \|h\nabla w\|)\|R\,w\|\,.
\end{aligned}$$

Combining the above bounds we get

$$\|h\nabla R\,w\|^2 \le C\left[\|h\nabla\,w\| + \|h\nabla R\,w\|\right]\mu\,\|R\,w\| + \|h\nabla\,w\|\,\|h\nabla R\,w\|\,.$$

The result then follows by applying the arithmetic-geometric mean inequality. □

Acknowledgment. The author would like to thank Georgios Akrivis, Manolis Georgoulis, Tristan Pryer and Andreas Veeser for many useful discussions and suggestions regarding this work.

REFERENCES

[1] I. BABUŠKA AND J. OSBORN, *Analysis of finite element methods for second order boundary value problems using mesh dependent norms*, Numer. Math., 34 (1980), pp. 41–62.

[2] I. BABUŠKA, J. OSBORN, AND J. PITKÄRANTA, *Analysis of mixed methods using mesh dependent norms*, Math. Comp., 35 (1980), pp. 1039–1062.

[3] S. C. BRENNER AND L. R. SCOTT, *The mathematical theory of finite element methods*, Springer-Verlag, New York, 1994.

[4] P. CIARLET AND J. LIONS, eds., *Finite Element Method (Part 1)*, vol. II of Handbook of Numerical Analysis, North Holland, Amsterdam, 1991.

[5] A. DEMLOW AND R. STEVENSON, *Convergence and quasi-optimality of an adaptive finite element method for controlling L_2 errors*, Numer. Math., 117 (2011), pp. 185–218.

[6] K. ERIKSSON, *An adaptive finite element method with efficient maximum norm error control for elliptic problems*, Math. Models Methods Appl. Sci., 4 (1994), pp. 313–329.

[7] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706–740.

[8] T. GUDI, *A new error analysis for discontinuous finite element methods for linear elliptic problems*, Math. Comp., 79 (2010), pp. 2169–2189.

[9] O. A. KARAKASHIAN AND C. G. MAKRIDAKIS, *Convergence of adaptive algorithms for elliptic problems*, tech. rep., In preparation, 2016.

[10] O. A. KARAKASHIAN AND F. PASCAL, *A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 2374–2399 (electronic).

[11] C. G. MAKRIDAKIS AND I. BABUŠKA, *On the stability of the discontinuous Galerkin method for the heat equation*, SIAM J. Numer. Anal., 34 (1997), pp. 389–401.

[12] P. MORIN, K. G. SIEBERT, AND A. VEESER, *Convergence of finite elements adapted for weaker norms*, in Applied and industrial mathematics in Italy II, vol. 75 of Ser. Adv. Math. Appl. Sci., World Sci. Publ., Hackensack, NJ, 2007, pp. 468–479.

[13] J. NITSCHE, *Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens*, Numer. Math., 11 (1968), pp. 346–348.

[14] J. A. NITSCHE AND A. H. SCHATZ, *Interior estimates for Ritz-Galerkin methods*, Math. Comp., 28 (1974), pp. 937–958.

[15] F. TANTARDINI, A. VEESER, AND R. VERFÜRTH, *Robust localization of the best error with finite elements in the reaction-diffusion norm*, Constr. Approx., 42 (2015), pp. 313–347.

[16] R. VERFÜRTH, *A posteriori error estimation techniques for finite element methods*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013.

MODELLING AND SC. COMPUTING, DMAM, UNIVERSITY OF CRETE / INSTITUTE FOR APPLIED AND COMPUTATIONAL MATHEMATICS, FORTH, GR 70013 HERAKLION, GREECE, AND MPS, UNIVERSITY OF SUSSEX, BRIGHTON BN1 9QH, UNITED KINGDOM

E-mail address: C.G.Makridakis@iacm.forth.gr